REMARKS — General

By the above amendment, the specification has been amended editorially and to clarify the invention patentably. The amended paragraphs in p. 6 and p. 10 describe the definition of the facial image enhancement module 200 in a more clear manner by giving an example of enhancing the user's facial feature image and giving the examples of the facial features in accordance with the embedded prior art FET system.

The example of superimposing an image of a pair of sunglasses onto the image of the user's eyes processed by the FET system in p. 6 and p. 10 is already shown as sunglasses image 108 in Fig. 1 and Fig. 5 of applicants' original application.

The definition of facial features, such as eyes, nose, and mouth, used in the applicants' proposed system is shown in the cited prior provisional patent application by R. Sharma and N. Jung, Method and System for Real-time Facial Image Enhancement, U.S. Provisional Patent. Application Number 60/394,324, July 8, 2002, (hereinafter Sharma and Jung), such as "Human facial features, the **smaller sub-objects on the face image**, can provide the useful local coordinate information within the face image in order to augment the human facial image", (lines 6-8, paragraph 1, page 4, Sharma and Jung)

Applicants have amended the claims to define the invention more particularly and distinctly so as to overcome the technical rejections and define the invention patentably over the prior art. Claims 1 to 4 have been integrated into the currently amended claim 6 to more particularly define the invention in a patentable manner over the cited prior art and make the currently amended claim 6 as an independent claim with the limitations in claims 1 to 4, as noted in the O.A.

Claims 5 has been rewritten as new claim 34 to more particularly define the invention in a patentable manner over the cited prior art.

Claims 7 and 8 have been rewritten into new claim 33 to more particularly define the invention in a patentable manner over the cited prior art.

Claim 11 has been integrated into the currently amended claim 16 to more particularly define the invention in a patentable manner over the cited prior art and make the currently amended claim 16 as an independent claim with the limitations in claim 11, as noted in the O.A.

Claims 20 have been rewritten as new claims 37 to more particularly define the invention in a patentable manner over the cited prior art.

Claims 21 to 30 have been rewritten into new claims 39 to 46 to more particularly define the invention in a patentable manner over the cited prior art, as noted in the O.A.

Claims 9, 10 and 17-19 have been amended as currently amended claims to more particularly define the invention in a patentable manner over the cited prior art.

Claims 12-15 have been cancelled.

The Rejection To The Claims Under 112

The last O.A. rejected claims 21—30 under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

The O.A. also noted that claim 21 recited the limitation "said user's selection for virtual object images on means for displaying output" in claim 21, step C, and there is insufficient antecedent basis for this limitation in the claim.

Applicants rewrote the claim 21 as new claims 39 to more particularly define the invention in a patentable manner and to have proper antecedent basis for the limitation in the claim, according to the O.A.

Applicants rewrote the claims 22 — 30 into the new claims 40 to 46 to incorporate the newly written claim 39 as dependent claims.

Accordingly applicants request reconsideration of this rejection.

The Rejection To The Claims Under 103

The O.A. rejected claims 1-5,7,11—15, and 20 under 35 U.S.C. 103(a) as being unpatentable over Lanier (6,400,374) in view of Hayama et al (US 2001/0034255) (hereinafter Hayama).

Applicants amended the claims as follows.

The Rejection Of Claim 1 and 11 on Lanier and Hayama Overcome

The O.A. rejected independent claim 1 on Lanier and Hayama. Claim 1 has been integrated into the currently amended claim 6 as a limitation of the newly made independent claim to define patentably over these references, and any combination thereof.

The O.A. rejected independent claim 11 on Lanier and Hayama. Claim 11 has been integrated into the currently amended claim 16 as a limitation of the newly made independent claim to define patentably over these references, and any combination thereof.

Applicants request reconsideration of these rejections, as now applicable to claim 6 and 16, for the following reasons:

(1) Fundamental approaches of facial image enhancement between the facial feature level process and the face level process are different. Facial feature level image processing deals with local facial components, such as eyes, nose, and mouth, on face images of users. Facial feature

detection is one example of facial feature level process in the art. However, face level processing, such as face detection, deals with the face image itself of the users. Thus, in the practiced art of facial image enhancement, these two approaches are treated differently and the details in their processing have their own different algorithms. A master's thesis by one of the applicants was entirely dedicated to a facial feature enhancement system, and it introduces a novel approach for the facial enhancement at the local facial feature level. Note that in a cited prior provisional patent application, Sharma and Jung discuss about this in detail. The facial image enhancement in Lanier and Hayama is done at the face level. Therefore, their approaches are fundamentally different approaches from the claim 6 and 16. The details of the difficulties for an efficient automatic, dynamic, and real-time processing facial feature enhancement are discussed later along with the cited prior art by Sharma and Jung, and it makes the applicants' proposed system unobvious over the prior art.

(2) The novel features of claim 6 and 16 produce new and unexpected results and hence are unobvious and patentable over these references. Applicants' novel approach for the facial enhancement at the local facial feature level in an uncontrolled environment for an audio-visual system creates a unique experience to the user and unobvious over the prior art.

(3) Even if Lanier and Hayama were to be combined in the manner proposed, the proposed combination would not show all the novel methods and physical features of claim 6 and 16, since neither of them mentions about the facial enhancement at the local facial feature level.

The above reasons also apply to the newly written independent claim 39, if it were to be rejected by the same reasons for the independent claim 1 and 11 in the last O.A.

The References and Differences of the Present Invention Thereof

Prior to discussing the claims and the above three points, applicants will first discuss the references and the general novelty of the present invention and its unobviousness over the references. The differences, in addition to the above three points, can be summarized as follows.

(1) Touch free interface is not introduced in either of the prior art. Touch-free interface in applicants' system enables a user to select virtual image objects on a display without physically touching any device in a novel way. Implementing a touch free interface is known to be significantly difficult task in the practiced art, computer vision image processing. For example, hand tracking and gesture recognition using camera in computer vision is one way to implement the touch free interface. Many researchers in the art have been studying how to come up with more robust and reliable hand-tracking algorithm. Due to this difficulty, the usage of touch free interface has not been found in many commercial products. Applicants use a commercially successful embodiment of touch free interface for the image object selection process and user interaction in a novel way. This embodiment is cited as a prior-art in reference, provisional patent application filed by R. Sharma, N. Krahnstoever, and E. Schapira, Method and System for Detecting Conscious Hand Movement Patterns and Computer-generated Visual Feedback for Facilitating Human-computer Interaction, U.S. Provisional Patent filed. April 2, 2002, (hereinafter Sharma and Krahnstoever). This novel usage of a touch-free interface in applicants' audio-visual system results in a unique experience to the user.

(2) The local facial feature enhancement in applicants' system is done in automatic, dynamic, and real-time processing fashion, which is by itself a difficult research problem. A master's thesis by one of the applicants was entirely dedicated to such facial feature enhancement system and the details of the difficulties for an automatic, dynamic, and real-time processing facial

feature enhancement can be found in the cited prior art by Sharma and Jung. For example, some exemplary challenges, which a real-time and automatic facial feature detection system has to overcome, are shown in (lines 10-28, paragraph 3, Sharma and Jung). The cited prior art FET system in Sharma and Jung suggests a novel approach, such as using a block processing in a mean crossing algorithm, to overcome the challenges, (lines 6-17, paragraph 2, Sharma and Jung). Applicants used the efficient automatic, dynamic, and real-time processing local facial feature enhancement system in a novel way in the proposed audio-visual entertainment system.

(3) Composition of virtual stage image with the enhanced local facial feature image is also novel feature in applicants' system, since neither of the prior art mentions a local facial feature level facial image enhancement.

(4) Although the prior art mentions about the background subtraction, applicants introduce a novel usage of the background subtraction technology. For example, applicants use the technology for providing fully immersive experience to the user by hiding the user's body image when the user is not moving, thus enticing the user to participate in a dance to see herself/himself on a display.

Lanier disclosed a system for superimposing a foreground image like a human head with face to the background image. However, Lanier's method does not handle the local facial features, such as eyes, nose, and mouth at all. Thus, granularity of facial image enhancement between Lanier and applicants is different. Lanier's method for detecting position of the user relies on a cumbersome position sensor, such as EM field receiver, (Lanier col. 2 lines 37-41, lines 50-53, col. 4 lines 8-10), whereas applicants use face detection for detecting the position of the user in the embedded FET system ("the FET system" 203 Fig. 2 and 3). Lanier's position sensor

approach is aimed to reduce the complexity in extracting a facial portion from a video in a typical digital image processing system (col. 5, lines 1-7), whereas applicants' approach is to use a novel digital image processing system, such as the FET system, for an efficient facial feature image detection and enhancement ("facial image enhancement module" 200 Fig. 2). An efficient facial feature image enhancement module in applicants' system makes it possible to run automatically, dynamically, and in real-time as described previously. Lanier uses background image which is tolerant of misalignments and misadjustments of the video image with the background image, since it is not always possible to adjust the size and placement of an image mask in his proposed system (col. 3, lines 21-25, second paragraph), whereas applicants use dynamic background (403 Fig. 4) which automatically adjust the size and orientation of the masked area. Applicant's boundary smoothing (602 Fig. 6) also makes the user look more immersive to the virtual state image. Lanier's method does not have specific information about the interface for the user interaction, whereas applicants' usage of a touch free interface (115 Fig. 2 and 3) for the audio-visual entertainment system is one of the novel features in the system. **Hayama** presents an image processing device which link first polygons to second polygons, but there is no description of facial feature level superimposition. Especially, in their definition description for "Polygons" (Hayama paragraph [0010] page 1), they do not specify about the local facial feature, such as eyes, nose, and mouth for their definition "part". Furthermore, it will be difficult to say, the local facial features, such as eyes, nose and mouth, are "of a body that moves as a single entity" in their definition for the "part", since the facial features are not moving entities. Even if their definition "part" were to include the facial features, such as eyes, nose, and mouth, they do not have any description how to detect these facial features from a face

image. Hayama does not have any description of interface with which to control the system nor does he describe about image object selection process, whereas applicants' usage of a touch free interface (115 Fig. 2 and 3) for the audio-visual entertainment system is one of the novel features in the system.

The last O.A. noted that the combination of Hayama's virtual state simulation technique and Lanier's audio-visual entertainment system could provide special effects and interesting video images in entertainment display system. However, neither Hayama nor Lanier mentions about the facial feature level processing and neither of them provides any information how actually this can be done. Their superimposition of images is limited by the granularity of face size rather than facial feature level, such as eyes, nose, and mouth as described above.

In these approaches, the human face image essentially becomes the superimposing object to the background templates or pre-handled video image sequences. However, applicants superimpose other virtual objects onto the human face image, thus further increasing the level of amusement. As mentioned in the application, "Human facial features can provide the useful local coordinate information within the face image in order to augment the human facial image" (lines 2-4, paragraph 1), and applying the local coordinate information within the face image to an audio-visual system is a distinctive feature over references. Applicants used a facial enhancement technology system, which not only detects the face and facial features efficiently, but also superimposes virtual objects on top of the user's face and facial features in real-time. This facial enhancement is another valuable feature addition to the audio-visual entertainment system along with the fully immersed body image into the dynamic virtual background. The superimposed objects move along with the user's arbitrary motion in real-time.

Neither Lanier nor Hayama mentions automatic, dynamic, and real-time facial feature detection and enhancement of facial images by the local facial feature image augmentation. Facial feature detection for the local facial feature is a known difficult task and doing it in real-time in an automatic and dynamic fashion is known to be more difficult task to the people in the art. Applicants use an efficient facial feature detection technology, and the usage of it to this context creates a novel experience to the users. The details of the FET system can be found in the cited prior provisional patent application by Sharma and Jung.

The virtual objects can be fanciful sunglasses, hat, hair wear, necklace, rings, beard, mustache, or anything else that can be attached to the human facial image. This whole process can transfigure the singer/dancer into a famous rock-star or celebrity on a stage and provides the user a new and exciting experience.

Neither Hayama nor Lanier mention about the role of user interaction and how the interaction can be accomplished. Applicant's touch free interaction paradigm allows a user to select image objects on a means for displaying in real-time, which significantly increases the novelty and amusement level of the audio-visual entertainment system. The user can change the virtual objects through a touch-free selection process. This process is achieved through tracking the user's hand motion in real-time.

Implementing a touch free interface is known to be significantly difficult task in the practiced art. For example, hand tracking and gesture recognition using camera in computer vision is one way to implement the touch free interface. Many researchers in the art have been studying how to come up with more robust and reliable hand-tracking algorithm. Due to this difficulty, the usage of touch free interface has not been found in many commercial products. Applicants use a

commercially known prior-art touch free interface in a new manner as cited in reference, provisional patent application filed by Sharma and Krahnstoever. The novel usage of this feature makes applicants' system unobvious over the prior art.

Neither Hayama nor Lanier describes fully immersive user integration with the background. Lanier used elliptical template for matching user image with the background image. In the prior art attempts for background subtraction, most approaches used a predefined static background or designated region, such as rectangular bounding box in a video loop. In the case of using a predefined static background, the background cannot be interactively controlled by the user in real-time. Although the user moves, the background image is not able to respond to the user's arbitrary motion. On the other hand, in the case of using the rectangular bounding box, although it might be possible to make the bounding box move along with the user's head motion, the user does not seem to appear to be fully immersed into the background image.

Unlike these previous attempts, our system, applicants use a dynamic background, which can change in real-time according to the user's arbitrary motion. The user's image appears to be fully immersed into the background, and the position of the user's image changes in any part of the background image as the user moves or dances while singing.

Another interesting feature of the dynamic background in the EVIKA system is that the user's image disappears behind the background if the user stands still. This adds an interesting and amusing value to the system, in which the user has to dance as long as the person wants to see himself on the screen. This feature can be utilized as a method to entice the user to participate in dancing. This also helps to encourage a group of users to participate.

The background can also be aesthetically augmented for decoration by the virtual objects. Virtual musical instrument images, such as guitars, pianos, and drums, can be added to the background. The individual instrument images can be attached to the user's image, and the instrument images can move along with the user's movement. The user can also play the virtual instrument by watching the instrument on screen and moving his hands around the position of the virtual instrument. This allows the user to participate further in the novel experience and therefore increases enjoyment.

The above mentioned novel and unobvious features in applicants' proposed system are foreign to Stelovsky(5,782,692), Corset(US 2002/0007718), and Nishitani et al(US 2003/0167908) (hereinafter Nishitani).

Stelovsky disclosed an interactive system for playing a game, educational or instructional sequence in conjunction with a prerecorded multimedia presentation consists at least motion video, sound and accompanying text, as noted in the O.A. However, Stelovsky's input/game controls are keyboard, mouse, joystick or other pointing device, voice command microphone, and speech digitizer, (col. 3, lines 4-7), which do not suggest a novel usage of a touch-free interface and selection process by the touch-free interface. Stelovsky does not show a facial image enhancement at the local facial feature level not the face level, a novel usage of a facial enhancement system, such as the FET, nor a novel usage of background and image composition, such as providing a fully immersive experience to the user and hiding effect when the user is not moving. Stelovsky does not show an automatic, dynamic, and real-time processing of facial feature enhancement.

Corset disclosed a karaoke system consisting of developing the concept of video immersion, as noted in the O.A. Corset's mixing and rendering device is a circuit using the shape information analyzed in the device for composing the user with a background pre-recorded video or audio-video delivered by a medium (col. 2, lines 6-9, paragraph [0010], page 1). However, Corset does not show a facial image enhancement at the local facial feature level not the face level, and a novel usage of a facial enhancement system, such as the FET. Corset's video mixing is not the composition of a virtual state image with the enhanced local facial feature image. Corset does not show an automatic, dynamic, and real-time processing of facial feature enhancement.

Nishitani disclosed an image with a guitar (Fig. 42). However, FIG. 42 is a diagram showing an example of another animation displayed during an automatic performance, and the guitar is not a virtual image. Applicants used a virtual guitar image 111 (Fig. 5 or lines 13-15, page 13), which can be added to the final virtual background image in the exemplary embodiment. Applicants' system allows a user to play the virtual musical instrument by pretending as if he or she actually plays the musical instrument while looking at the musical instrument image on the display. Nishitani does not show a facial image enhancement at the local facial feature level not the face level, a novel usage of a facial enhancement system, such as the FET, nor a novel usage of background and image composition, such as providing a fully immersive experience to the user and hiding effect when the user is not moving. Nishitani does not show an automatic, dynamic, and real-time processing of facial feature enhancement.

Overall, the above-mentioned prior arts do not show the novel features in applicants' proposed system, which can comprise the followings.

(1) Facial image enhancement at the local facial feature level not the face level, and a novel usage of a facial enhancement system, such as the FET.

(2) Composition of virtual state image with the enhanced local facial feature image.

(3) Usage of arbitrary background for the processing of virtual state image composition, enhancement of local facial feature image, and interaction by a touch-free interface.

(4) Automatic, dynamic, and real-time processing of facial feature enhancement.

(5) Using a commercially successful embodiment of touch free interface for the image object selection process and user interaction.

(6) Novel usage of background subtraction, such as providing a fully immersive experience to the user and hiding effect when the user is not moving.

The Rejection Of Claim 8 on Lanier, Hayama, and Kumar Overcome

The O.A. rejected dependent claim 8 on Lanier, Hayama, and Kumar et al (6,692,259) (hereinafter Kumar). Claim 8 has been integrated into the new claim 33 as a limitation to define patentably over these references, and any combination thereof. The general novelty of the present invention and its unobviousness over Lanier and Hayama described above applies to claim 8. In addition to this, applicants will discuss Kumar in the following paragraphs.

Kumar requires an initiation process, which is typically accomplished by using a remote control, col. 5, lines 14-21. Applicants use touch free interface to initiate the music selection. Kumar describes his background subtraction technologies permit the user to interact with the Karaoke system. Applicants use touch free interface technologies to enable the selection process by the user. One particular embodiment of the touch free interface is shown in the cited prior provisional patent application filed by Sharma and Krahnstoever. Kumar's gesture analysis is

contributed to the composition of images, col. 5, lines 41-44. However, Kumar does not specifically mention how the gesture analysis can be used for selection process. Applicant's touch free interaction paradigm allows a user to select image objects on a means for displaying in real-time directly not as a by-product. The user can change the virtual objects through a touch-free selection process. This process is achieved through tracking the user's hand motion in real-time.

The Dependent Claims Are a Fortiori Patentable Over Lanier and Hayama

The currently amended claims 9 and 10 incorporate all the subject matter of claim 6 and add additional subject matter, which makes them a fortiori and independently patentable over these references.

New dependent claims 31 to 34 incorporate all the subject matter of claim 6 and add additional subject matter, which makes them a fortiori and independently patentable over these references.

Claim 9 further adds a step for using said user's movement to trigger said dynamically changing virtual background images, whereby without said user's movement, said user's image could disappear behind said virtual background image. Neither Lanier nor Hayama do this. Applicants' feature adds an interesting and amusing value to the system, in which said user has to dance as long as said user wants to see herself/himself on a means for displaying output, and this feature can be utilized as a method for said user to participate in a dance in front of said audio-visual entertainment system.

Claim 10 further adds a step for attaching musical instrument images, such as a guitar image or a violin image, to said user's body image. Neither Lanier nor Hayama do this. The attached musical instrument images dynamically move along with said user's arbitrary motion in real-

time, and the user can also play the musical instrument by pretending as if he or she actually plays the musical instrument while looking at said musical instrument image on said means for displaying output.

Claim 31 further adds a step for processing the facial image enhancement automatically, dynamically, and in real-time. Neither Lanier nor Hayama mentions automatic, dynamic, and real-time facial feature detection and enhancement of facial images by the local facial feature image augmentation.

Claim 32 further adds a touch free interface for processing virtual object image selection and processing music selection. Neither Lanier nor Hayama has the touch free interface.

Claim 33 further adds a step for processing said virtual object image selection and music selection by said touch free interface and the composition of the virtual stage images on any arbitrary background in the actual environment rather than a controlled background, such as a blue-screen style background, whereby the dynamic background construction can be processed by an adaptive background subtraction algorithm. Neither Lanier nor Hayama use the arbitrary background in the actual environment for the selection process by a touch free interface.

Claim 34 further adds a step for combining the enhanced facial images of said user or said plurality of users and body images of said user or said plurality of users with said dynamically changing virtual background images. This is entirely foreign to Lanier and Hayama, or any combination thereof, since the systems in the references do not process the facial features, such as eyes, nose, or mouth, but face image itself at a larger image resolution level.

Claim 17 further adds means for using a facial image enhancement process. Neither Lanier nor Hayama process the facial image enhancement at the level of local facial features, such as eyes, nose, and mouth.

Claim 18 further adds means for using the embedded FET system for said facial image enhancement process. Neither Lanier nor Hayama have the embedded FET system.

Claim 19 further adds means for preparing said virtual object images, such as musical instrument images and stage images, off-line. Neither Lanier nor Hayama mention about this.

Claim 35 further adds means for enhancing the facial images automatically, dynamically, and in real-time. Neither Lanier nor Hayama mentions automatic, dynamic, and real-time facial feature detection and enhancement of facial images by the local facial feature image augmentation.

Claim 36 further adds means for processing virtual object image selection, processing music selection, and composing virtual stage images, wherein the selection is processed by a touch free interface. Neither Lanier nor Hayama do this.

Claim 37 further adds means for processing any arbitrary background in the actual environment rather than a controlled background, such as a blue-screen style background, not only for constructing dynamically changing virtual background images but also for processing user interaction by a touch-free interface, whereby said dynamically changing virtual background images are background images which change according to arbitrary movement of said user or said plurality of users in real-time. Neither Lanier nor Hayama mentions about a touch-free interface in an uncontrolled background. As stated above, it is a novel usage to apply the digital image processing based touch-free interface for this purpose in an audio-visual entertainment system.

Claim 38 further adds means for combining the enhanced facial images of said user or said plurality of users and body images of said user or said plurality of users with said dynamically changing virtual background images. Neither Lanier nor Hayama mentions composition of virtual background images with enhanced facial images, which are accomplished at the local facial feature level, such as eyes, nose, and mouth.

Accordingly applicants submit that the dependent claims are a fortiori patentable and should also be allowed.

The Objection To The Claims

Claims 6, 9 – 10, and 16 – 19 were objected to as being dependent upon a rejected base claim.

Applicants rewrote the claim 6 in independent form including all of the limitations of the previous base claim, according to the O.A.

Applicants rewrote the claims 9—10 to be dependent on the rewritten independent claim 6, which includes all of the limitations of the previous base claim, according to the O.A.

Applicants rewrote the claim 16 in independent form including all of the limitations of the previous base claim, according to the O.A.

Applicants rewrote the claims 17—19 to be dependent on the rewritten independent claim 16, which includes all of the limitations of the previous base claim, according to the O.A.

CONCLUSION

For all the above reasons, applicants submit that the specification and claims are now in proper form, and that the claims all define patentably over the prior art. Therefore they submit that this application is not in condition for allowance, which action they respectfully solicit.

Conditional Request for Constructive Assistance

Applicants have amended the specification and claims of this application so that they are proper, definite, and define novel structure, which is also unobvious. If, for any reason this application is not believed to be in full condition for allowance, applicants **very respectfully request** the constructive assistance and suggestions of the Examiner pursuant to M.P.E.P. 2173.02 and 707.07(j) in order that the undersigned can place this application in allowable condition.

Very respectfully,

Namsoon Jung Rajeev Sharma

----------Applicants Pro Se----------

403 South Allen Street, Suite 104

State College, PA 16801

Tel. (814) 867-8977; Fax (814) 867-8957

Certificate of Mailing

I hereby certify that this correspondence, and attachments, if any, will be deposited with the United States Postal Service by First Class Mail, postage prepaid, in an envelope addressed to "Box Non-Fee Amendments, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450" on the date below.

Date: 5|20|2005 Inventor's Signature: Ishaque